Simulating Gaussian Stationary Processes with Unbounded Spectra

Brandon Whitcher
EURANDOM
P.O. Box 513
5600 MB Eindhoven
The Netherlands
whitcher@eurandom.tue.nl

October 1, 1999

Abstract

We propose a new method for simulating a Gaussian process, whose spectrum diverges at one frequency in $[0, \frac{1}{2}]$ (not necessarily at zero). The method utilizes a generalization of the discrete wavelet transform, the discrete wavelet packet transform (DWPT), and only requires explicit knowledge of the spectral density function of the process – not its autocovariance sequence. An orthonormal basis is selected such that the spectrum of the wavelet coefficients is as shallow as possible, thus producing approximately uncorrelated wavelet coefficients. We compare this method to a popular time-domain technique based on the Levinson-Durbin recursions. Simulations show that the DWPT-based method performs comparably to the time-domain technique for a variety of sample sizes and processes – at significantly reduced computational time. The degree of approximation and reduction in computer time may be adjusted through selection of the orthonormal basis.

Some key words: Autocovariance, Discrete wavelet packet transform; Gegenbauer process; Orthonormal basis; Seasonal persistent process; Time series.

1 Introduction

The use of fractional ARIMAs as models for time series exhibiting long-range dependence is now quite widespread. Such time series are characterized by an autocovariance sequence $\{s_\tau\}$ which diverges; i.e., $\sum_{\tau=0}^{\infty} s_\tau = \infty$. Let $\{X_t\}$ be a stochastic process whose dth order backward difference

$$(1-B)^d X_t = \epsilon_t \tag{1}$$

is a stationary process, where $-\frac{1}{2} < d < \frac{1}{2}$ and B is the backward difference operator. This process is stationary and invertible. If $\{\epsilon_t\}$ is a Gaussian white noise process with variance σ_ϵ^2, then $\{X_t\}$ is the simplest case of a fractional ARIMA process, a fractional ARIMA$(0, d, 0)$, which we refer to as a fractional difference process. The spectral density function (SDF) of $\{X_t\}$ is

$$S_X(f) \equiv \sigma_\epsilon^2 |2\sin(\pi f)|^{-2d} \quad \text{for} \quad -\frac{1}{2} < f < \frac{1}{2},$$

so that $S_X(f) \to \infty$ as $f \to 0$ and, thus, the SDF diverges at zero frequency. Further introductions to fractional difference and related processes can be found in, e.g., Granger and Joyeux (1980) and Hosking (1981).

Both time- and frequency domain techniques have been established for the simulation of such long memory processes (see Percival (1992) and references therein), the partitioning of the time-frequency plane by the discrete wavelet transform (DWT) makes it a natural alternative to the discrete Fourier transform. This has been investigated by, e.g., Wornell (1993), Masry (1993) and McCoy and Walden (1996). The wavelet coefficients of a long memory process are approximately uncorrelated. Hence, simulation may be performed by simply generating Gaussian random variables with zero mean and variance proportional to the integrated spectrum over octave bands. Applying the inverse DWT produces a realization of the long memory process.

The time-domain technique is readily adaptable to processes whose spectral energy is not dominated in the lower frequencies (e.g., an MA(1) process with $\theta \to -1$), as long as the autocovariance

sequence is known. Percival (1992) extended the frequency-domain technique to non-stationary power law processes, but these still require the spectrum to be dominated by low-frequency energy. The DWT simply cannot adapt and we must use the discrete wavelet packet transform (DWPT), which creates a redundant set of wavelet coefficients at each level of the transform (Wickerhauser 1994, Ch. 7). To determine an appropriate orthonormal basis, we propose a method based on achieving the most shallow spectrum for each level of the DWPT, thus producing the least correlated wavelet coefficients for a much wider class of stochastic processes. Thus, simulation may be performed through a method similar to that of long-memory processes. The technique proposed by McCoy and Walden (1996) may be considered a special case of this method.

In Section 2 we introduce the seasonal persistent process as an example of a time series with unbounded spectra. Because such processes do not possess a closed-form expression for their autocovariance function, we investigate the convergence of an asymptotic approximation. The DWPT is briefly defined in Section 3 and the basis selection technique is illustrated. Simulation results are presented in Section 4 where the DWPT method is compared to a time-domain technique (utilizing the Levinson-Durbin recursions). Conclusions are presented and recommendations given at the end.

2 Seasonal Persistent Processes

A simple generalization of the model given in (1) was mentioned, in passing, by Hosking (1981) and allows the singularity in the spectrum to be located at any frequency $0 \leq f \leq \frac{1}{2}$. Such a process has been referred to as a Gegenbauer process (Gray *et al.* 1989) and also a seasonal persistent process (SPP) (Anděl 1986). We prefer the latter term because it more accurately and concisely describes the content of the time series. That is, a sinusoid of particular frequency is associated with the singularity present in the spectral density function (SDF) thus causing a persistent oscillation in

the process. Gray *et al.* (1989) fit a seasonal long-memory model to the Wolfer sunspot data, where short-range dependence was allowed through fitting ARMA components. Recent attention has also appeared in the economics literature, where Ooms (1995) fit seasonal long-memory models to the U.S. gross national product and a time series of Danish shipping records. Arteche and Robinson (1999) discuss various models for seasonal long memory and propose a semi-parametric estimation procedure.

Let $\{Y_t\}$ be a stochastic process such that

$$\left(1 - 2\phi B + B^2\right)^{\delta} Y_t = \epsilon_t \tag{2}$$

is a stationary process, then $\{Y_t\}$ is an SPP. Gray *et al.* (1989) showed that $\{Y_t\}$ is stationary and invertible for $|\phi| = 1$ and $-\frac{1}{4} < \delta < \frac{1}{4}$ or $|\phi| < 1$ and $-\frac{1}{2} < \delta < \frac{1}{2}$.

Clearly, the definition of an SPP also includes a fractional difference process. When $\phi = 1$ we have that $\{Y_t\}$ is a fractional difference process given by (1) with fractional difference parameter $d = 2\delta$. If $\{\epsilon_t\}$ is a Gaussian white noise process, then $\{Y_t\}$ is also called a Gegenbauer process since (2) may be written as

$$Y_t = \sum_{k=0}^{\infty} C_{k,\phi}^{(\delta)} \epsilon_{t-k}, \tag{3}$$

where $C_{k,\phi}^{(\delta)}$ is a Gegenbauer polynomial (Rainville 1960, Ch. 17). The SDF of $\{Y_t\}$ is given by

$$S_Y(f) = \sigma_\epsilon^2 \{2|\cos(2\pi f) - \phi|\}^{-2\delta}, \quad \text{for} \quad -\frac{1}{2} < f < \frac{1}{2}, \tag{4}$$

so that $S_Y(f)$ becomes unbounded at frequency $f_0 \equiv (\cos^{-1}\phi)/(2\pi)$, sometimes called the Gegenbauer frequency.

The autocovariance sequence of an SPP may be expressed via

$$\gamma_\tau = \int_{-1/2}^{1/2} S_Y(f) \cos(2\pi f \tau)\, df. \tag{5}$$

An explicit solution is known only for special cases (Anděl 1986). Gray *et al.* (1994) showed that the autocorrelation sequence of an SPP is given by

$$\rho_\tau \sim \tau^{2\delta-1}\cos(2\pi f_0 \tau) \quad \text{as} \quad \tau \to \infty. \tag{6}$$

Two sequences are related via $\alpha_\tau \sim \beta_\tau$ as $\tau \to \infty$ if $\lim_{\tau\to\infty}\{\alpha_\tau/\beta_\tau\} = c$ where c is a finite nonzero constant. The result in (6) was given in Hosking (1982) without proof.

In the time-domain method of simulation, an approximation to the integral would greatly reduce the computational burden. Figure 1 displays the asymptotic approximation (6) for various SPPs defined in Figures 8-11 from Anděl (1986). Note, there is an error in the caption of Figure 11 in Anděl (1986), the frequency is given as $\omega = 2\pi f_0 = 0.1$, but ϕ is incorrectly stated as 0.955 whereas it should be 0.995. The numeric integration of (5) was performed via a FORTRAN routine from Ford (1991). When the fractional difference parameter δ is large, the asymptotic approximation matches the autocorrelation sequence from numeric integration well – only slightly over-estimating it. This apparent over-estimation is quite pronounced when δ is relatively small, and persists for a large number of lags. We will utilize this approximation for lags of 100 or greater.

3 A Discrete Wavelet Packet Transform Method

3.1 The Discrete Wavelet Packet Transform

The orthonormal discrete wavelet transform (DWT) is known to approximately decorrelate long memory processes (Tewfik and Kim 1992; Wornell 1996). It does this through band-pass filtering the process in such a way that the spectrum in each pass band is approximately constant. Using this property, McCoy and Walden (1996) outlined a procedure for simulating a fractional difference process using the DWT where only the overall variance and the band-pass variances of the process are required. Explicit knowledge of the covariance structure is not necessary in order to simulate.

The DWT has a very specific band-pass structure which partitions the spectrum of a long

memory process finer and finer as $f \to 0$ – where the spectrum is unbounded. This is done through a succession of filtering and downsampling operations; see, e.g., Percival and Walden (1999, Ch. 4) for an introduction to the DWT. In order to exploit the approximate decorrelation property for SPPs we need to generalize the partitioning scheme of the DWT. This is easily obtained by performing the discrete wavelet packet transform (DWPT) on the process; see, e.g., Wickerhauser (1994, Ch. 7) and Percival and Walden (1999, Ch. 6). Instead of one particular filtering sequence, the DWPT executes all possible filtering combinations to obtain a wavelet packet tree, denoted by $\mathcal{T}$. An orthonormal basis $\mathcal{B} \subset \mathcal{T}$ is obtained when a collection of DWPT coefficients is chosen, whose ideal band-pass frequencies are disjoint and cover $[0, \frac{1}{2}]$.

Let $h_0, \ldots, h_{L-1}$ be the unit scale wavelet (high-pass) filter coefficients from a Daubechies compactly supported wavelet family (Daubechies 1992) of even length L. In the future, we will denote the Daubechies family of extremal phase compactly supported wavelets with D(L) and the Daubechies family of least asymmetric compactly supported wavelets with LA(L). The scaling (low-pass) coefficients may be computed via the quadrature mirror relationship

$$g_l = (-1)^{l+1} h_{L-l-1}, \quad l = 0, \ldots, L-1.$$

Now define

$$u_{n,l} \equiv \begin{cases} g_l, & \text{if } n \bmod 4 = 0 \text{ or } 3; \\ h_l, & \text{if } n \bmod 4 = 1 \text{ or } 2, \end{cases}$$

to be the appropriate filter at a given node of the wavelet packet tree.

Let $\mathbf{X}$ be a length N vector of observations and $\mathbf{W}_{j,n}$ denote the vector of wavelet coefficients associated with the frequency interval $\lambda_{j,n} \equiv \left(\frac{n}{2^{j+1}}, \frac{(n+1)}{2^{j+1}} \right]$. Let $W_{j,n,t}$ denote the tth element of the length $N_j \equiv N/2^j$ vector $\mathbf{W}_{j,n}$. Given the vector of DWPT coefficients $\mathbf{W}_{j-1,\lfloor \frac{n}{2} \rfloor}$, we compute

$W_{j,n,t}$ via

$$W_{j,n,t} \equiv \sum_{l=0}^{L-1} u_{n,l} W_{j-1,\lfloor \frac{n}{2} \rfloor, 2t+1-l \bmod N_{j-1}}, \quad t = 0, 1, \ldots, N_j - 1,$$

where $L_j = (2^j - 1)(L - 1) + 1$ is the length of a level j wavelet filter. To start the recursion set $\mathbf{W}_{0,0} = \mathbf{X}$. This is only one possible formulation of the DWPT, we may also directly filter the observations by generating unique filter coefficients at each level or apply a series of matrix operations; see Percival and Walden (1999, Ch. 6) for more details on this and other formulations. As with the DWT, the DWPT is most efficiently computed using a pyramid algorithm. The algorithm has $O(N \log N)$ operations, like the fast Fourier transform.

3.2 Selecting the Basis

Let $\mathcal{T} = \{(j, n) \mid j = 0, \ldots, J; n = 0, \ldots, 2^j - 1\}$ be the collection of all doublets (j, n) which form the indices of the nodes of a wavelet packet tree. Let $\mathcal{B}$ be a collection of doublets (j, n) which correspond with an orthonormal basis. Although several methods exist to chose a particular collection of nodes (see Chen *et al.* (1999) for an overview), a simple one is as follows. After the initial high- and low-pass filtering operation, simply apply a successive high- and low-pass filtering operation to the portion of the spectrum which contains the Gegenbauer frequency. The idea behind this is that we want to obtain band-pass spectra which are as flat as possible, hence finer partitioning of the frequency axis is necessary where the spectrum is steepest. Figure 2b shows this ideal basis in terms of the time-scale plane down to level $J = 6$ for an SPP with SDF shown in Figure 2a; i.e.,

$$\mathcal{B} = \{(1, 1), (2, 1), (3, 0), (4, 3), (5, 4), (6, 10), (6, 11)\}.$$

When synthesizing fractional difference processes, as in McCoy and Walden (1996), the basis associated with the DWT is sufficient for producing approximately uncorrelated wavelet coefficients

at each scale. One inherent property of any wavelet filter is zero mean, which corresponds to its squared gain function having value 0 at frequency 0 (in fact, the Daubechies families of wavelet filters considered here have $L/2$ vanishing moments). This property is key to producing approximately uncorrelated wavelet coefficients for so-called long memory processes and is distinctly lacking when the asymptote is allowed to vary throughout frequencies $f_0 \in [0, \frac{1}{2}]$. The immediate consequence is that the ideal basis in Figure 2b will not be sufficient to guarantee approximately uncorrelated wavelet coefficients for an arbitrary wavelet filter.

Using the SPP described in Figure 2a, the SDF of $\mathbf{W}_{1,1}$ is given for three different Daubechies wavelet filters of length $L \in \{2, 4, 8\}$ in Figure 3. Even though the ideal pass-band is $\frac{1}{4} \leq |f| \leq \frac{1}{2}$, and $f_0 = \frac{1}{12}$, the poor approximation to an ideal band-pass filter is apparent for the Haar and D(4) wavelet filters. They are insufficient for producing approximately uncorrelated wavelet coefficients for this particular process. The LA(8) wavelet filter allows only a very small spike of energy from the singularity in the spectrum of the SPP.

One way to overcome the poor approximation of the wavelet filters to that of ideal band-pass filters would be to select a basis where the squared gain function (modulus squared of the DFT) of the wavelet filter associated with $\mathbf{W}_{j,n}$ is sufficiently small at the Gegenbauer frequency. Let us define $\mathcal{U}_{j,n}(f_0) \equiv |U_{j,n}(f_0)|^2$, where $U_{j,n}(f)$ is the DFT of

$$u_{j,n,l} \equiv \sum_{k=0}^{L-1} u_{n,k} u_{j-1,\lfloor \frac{n}{2} \rfloor, l-2^{j-1}k}, \quad l = 0, \ldots, L_j - 1,$$

with $u_{1,0,l} \equiv g_l$ and $u_{1,1,l} \equiv h_l$ (Percival and Walden 1999, Ch. 6). The partition of the time-scale plane would therefore depend upon the choice of wavelet filter, overall depth of the partition being inversely proportional to the length of the wavelet filter. Hence, the basis selection procedure involves selecting the combination of wavelet bases which achieve $\mathcal{U}_{j,n}(f_0) < \varepsilon$, for some $\varepsilon > 0$, at the minimum level j.

Figure 4 gives the 'best' basis, according to the criterion $\mathcal{U}_{j,n}(f_0) < 0.01$, for various Daubechies wavelet filters $L \in \{2, 4, 8, 16\}$ applied to the SPP shown in Figure 2a. As to be expected, the shorter wavelet filters $L \in \{2, 4\}$ are poor approximations to an ideal band-pass filter and produce an orthonormal basis which is nothing like the ideal basis shown in Figure 2b. The Daubechies wavelet filter with $L = 8$ produces a basis that follows the general shape of the ideal one, but suffers around the Gegenbauer frequency. The longest wavelet filter, and therefore best approximation to an ideal band-pass filter, generates a basis function similar to the ideal basis.

3.3 Simulation via the DWPT

The bandpass variance $B_{j,n}$ for an SPP, with SDF given in (4), in the frequency interval $-\frac{n}{2^{j+1}} \leq |f| \leq \frac{n+1}{2^{j+1}}$ is

$$B_{j,n} \equiv 2 \int_{\frac{n}{2^{j+1}}}^{\frac{n+1}{2^{j+1}}} \sigma_\epsilon^2 \{2|\cos(2\pi f) - \phi|\}^{-2\delta} \, df = 2 \cdot 4^{-\delta} \sigma_\epsilon^2 \int_{\frac{n}{2^{j+1}}}^{\frac{n+1}{2^{j+1}}} |\cos(2\pi f) - \phi|^{-2\delta} \, df. \tag{7}$$

As in McCoy and Walden (1996), we replace the true SDF at each frequency band with a constant $S_{j,n} = S_{j,n}(f)$, for all f, such that the band-pass variances are equal. This step assumes the SDF is slowly varying across the frequency interval $\lambda_{j,n}$ (c.f. Section 3.2). Integrating the constant SDF over $\lambda_{j,n}$ gives

$$\int_{\frac{n}{2^{j+1}}}^{\frac{n+1}{2^{j+1}}} S_{j,n} \, df = S_{j,n} 2^{-j-1}.$$

Equating this to the band-pass variance gives

$$2 S_{j,n} 2^{-j-1} = B_{j,n} \quad \Longrightarrow \quad S_{j,n} = 2^j B_{j,n}. \tag{8}$$

The variance of $W_{j,n,t}$ is therefore given by $S_{j,n}$ because of the band-pass nature of the transform.

Let us consider simulating a length $N = 2^J$ SPP with known parameters ϕ and δ. The DWPT-based simulation procedure can be implemented as follows:

1. We must determine the appropriate orthonormal transform $\mathcal{B}$ from the wavelet packet tree. This is done by computing, for levels $j = 1, \ldots, J-1$, the squared gain function of the desired wavelet filter which satisfies $\mathcal{U}_{j,n}(f_0) < \varepsilon$ to be included in $\mathcal{B}$ (we used $\varepsilon = 0.01$). This corresponds with the idea of partitioning the frequencies where the spectrum is steepest. At the final level J, all remaining nodes for uncovered frequency intervals must be included in order to complete the orthonormal basis.

2. We need to calculate the band-pass variances $B_{j,n}$, $(j,n) \in \mathcal{B}$. They are defined in (7) and obtained via numeric integration. We use (8) to obtain $S_{j,n}$. Each DWPT coefficient $W_{j,n,t}$, $(j,n) \in \mathcal{B}$, $t = 1, \ldots, 2^{J-j}$ is a independent Gaussian random variable with zero mean and variance $S_{j,n}$.

3. Once the DWPT coefficients have been generated we may organize them and apply the inverse DWPT defined by $\mathcal{B}$ to obtain the simulated SPP $\widetilde{Y}_1, \ldots, \widetilde{Y}_N$.

The numeric integration for $j = 1, \ldots, J-1$ required in [2] is easily computed using the routine QAWO from QUADPACK (Piessons *et al.* 1983), since the SDF is bounded on the interval. At the final level of the DWPT, the SDF of $\mathbf{W}_{J,n}$, such that $f_0 \in \lambda_{J,n}$, will not be bounded. The numeric integration may still be performed by splitting the integral at the Gegenbauer frequency.

4 Simulation Results

Gray *et al.* (1989) used (3) to simulate an SPP by truncating the infinite sum to 290,000 terms. We abandon this approach for two reasons, it is very computer intensive and it depends on the convergence of the process. Woodward *et al.* (1998) instead used a time domain approach that utilizes the Levinson–Durbin recursions; see, e.g., Hosking (1984) and Percival (1992). The auto-covariance sequence is required for this procedure and was calculated via numeric integration. We

call this the *Hosking method*.

Figure 5 shows the results of a small simulation study (500 realizations) demonstrating the ability of the Hosking method to generate SPPs with parameters given in Figure 1. As to be expected, the Hosking method produces realizations with excellent second-order properties. The sample ACVS, averaged over the 500 simulations, follows the true ACVS very well for all four processes. The 5% and 95% points are provided to indicate variability of the sample ACVS.

The DWPT-based method outlined in Section 3.3 was performed to simulate SPPs with parameters identical to those in Figure 5. The basis used was $\mathcal{B}_7 \equiv \{(7,n) \,|\, n = 0, \ldots, 63\}$ and the wavelet filter was the minimum-bandwidth discrete-time wavelet with 16 nonzero coefficients. We refer to this class of wavelet filters as MB(L); see Morris and Peravali (1999) for details on this family of wavelets. The minimum-bandwidth discrete-time wavelets exhibit better frequency resolution than Daubechies families of wavelets at comparable lengths. Looking at Figure 6, the mean ACVS follows the true ACVS almost as well as the Hosking method and with reduced variability as deomonstrated by the 5% and 95% points. The DWPT method encounters difficulty with the final SPP, $\delta = 0.3$ and $\phi = 0.995$, at all but the smallest sample size.

Although graphically comparing the two methods is illuminating, a quantitative comparison between the true ACVS and the averaged sample ACVS is provided by the one-sided least squares difference (Percival 1992)

$$\mathrm{LSD}(M) \equiv \sum_{\tau=0}^{M} |\hat{s}_\tau - s_\tau|^2 ,$$

where $\{\hat{s}_\tau\}$ is the averaged ACVS and $\{s_\tau\}$ is the true ACVS. We prefer to use only lags up to $M = 100$ versus all possible lags. Table 1 provides LSD(100) from a more extensive simulation study comparing the Hosking and DWPT methods under a variety of conditions. The basis used for the DWPT method was $\mathcal{B}_7 \equiv \{(7,n) \,|\, n = 0, \ldots, 63\}$ for all wavelet filters. The Hosking method

improves dramatically as the sample size increases for all four processes, remember only the first 100 lags were used to compute LSD(100). The DWPT method, when utilizing the MB(8), D(16) and LA(16) wavelet filters, performs similarly to one another and slightly less accurate than the Hosking method. This is because the squared gain functions are identical for the D(16) and LA(16) wavelets and very close to the MB(8) wavelet. When using the MB(16) wavelet filter, the DWPT method performs better than the other wavelet filters and comparably to the Hosking method. No noticeable improvement was gained by using the MB(24) wavelet filter.

Table 2 provides LSD(100) from another simulation study, where the basis used for the DWPT method was computed using the criterion $\mathcal{U}_{j,n}(f_0) < 0.01$. The first two SPPs, with $\delta = 0.3$ and $\phi = 0.866$, have the same basis since $f_0 = \frac{1}{12}$; see Figure 4. The SPP, with $\delta = 0.3$ and $\phi = -0.6$, has the asymptote around $f_0 = 0.352$ and looks like a reflected version of Figure 4. The SPP, with $\delta = 0.3$ and $\phi = 0.995$ has its asymptote around $f_0 = 0.016$, quite close to zero. Hence, its basis looks very close to the DWT basis used for long memory processes, with differences appearing at the larger scales and lower frequencies only. The average squared differences between the true ACVS and sample ACVS are very similar to those obtained using the basis $\mathcal{B}_7$ as given in Table 1. Hence, the method proposed in Section 3.2 is generating an adaptive orthonormal basis that produces approximately uncorrelated wavelet coefficients. This is a much more parsimonious basis than simply using all nodes from a particular level of the transform.

Computational efficiency of an algorithm is an important issue in simulation methodology. The DWPT has $O(N \log N)$ computational complexity, while the Levinson-Durbin recursions require $O(N^2)$ operations. Both methods were implemented in S-Plus with the intensive computations being written in C. Using the function `unix.time()`, Figure 7 compares the computational time needed to simulate SPPs for a variety of sample sizes on an SGI workstation. It appears the DWPT-based method requires a minimum of a half-second in order to compute one realization of

small to moderate length. After $N \geq 2^{10} = 1024$, the computation time increases at a relatively slow rate and agrees with a similar technique for simulating long-range dependent network traffic (Ribeiro *et al.* 1999). The Hosking method is quite fast for small sample sizes, is comparable around $N = 512$, then greatly exceeds the DWPT method for larger sample sizes. In fact, S-Plus failed to successfully execute when $N \geq 2^{12}$ – subsequent simulation times are given by extrapolation. In order to compute the Hosking method for large sample sizes, efficient storage and computation of the partial autocorrelations ($N \times N$ matrix) would be required.

5 Discussion

A new method has been proposed in order to simulate Gaussian stationary processes, whose SDF is unbounded at a specific frequency $0 \leq f \leq \frac{1}{2}$, through the discrete wavelet packet transform. While not an 'exact' method, as compared with the procedure given by Hosking (1984), it performs comparably in Monte Carlo simulations. Given the relative efficiency of the DWPT, it is faster to compute than the Hosking method and is 'adaptive' in the sense that a variety of orthonormal bases may be selected – giving the user a choice between precision and computational speed.

An obvious extension of this method would be to allow multiple singularities in the SDF of the process. Some examples of such models are the k-factor GARMA process (Woodward *et al.* 1998) and ARUMA process (Giraitis and Leipus 1995). The only modification in the proposed method would be to select an orthonormal basis $\mathcal{B}$ which takes into account these multiple singularities. This involves evaluating the squared gain function of the wavelet filter and selecting those basis functions which satisfy a user-defined threshold for all frequencies with associated singularities.

The DWPT method for simulating SPPs gives a new perspective on determining the "appropriate" basis for the transform. When using the DWPT to analyze (decompose) time series one may select the basis by, e.g., the best basis algorithm (Coifman and Wickerhauser 1992) or match-

ing pursuit (Mallat and Zhang 1993). In this article we were interested in the synthesis of time series using a different type of basis – one that minimizes the correlation structure of the DWPT coefficients. In the analysis of SPPs, we have a best basis by knowing the Gegenbauer frequency. If we were given a series with an unknown Gegenbauer frequency, we could apply the DWPT and then test the DWPT coefficients using a standard test for white noise (e.g., the cumulative periodogram test). A basis would then be selected which produces DWPT coefficients with least residual correlation at the lowest level possible.

References

Anděl, J. (1986). Long memory time series models. *Kybernetika 22*(2), 105–123.

Arteche, J. and P. M. Robinson (1999). Seasonal and cyclical long memory. In S. Ghosh (Ed.), *Asymptotics, Nonparametrics, and Time Series*, Volume 128 of *STATISTICS: Textbooks and Monographs*, pp. 115–148. New York: Marcel Dekker.

Chen, S. S. B., D. L. Donoho, and M. A. Saunders (1999). Atomic decomposition by basis pursuit. *SIAM Journal of Scientific Computing 20*(1), 33–61.

Coifman, R. R. and M. V. Wickerhauser (1992). Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory 38*(2), 713–718.

Daubechies, I. (1992). *Ten Lectures on Wavelets*, Volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia: Society for Industrial and Applied Mathematics.

Ford, C. R. (1991). *The Gegenbauer and Gegenbauer Autoregressive Moving Average Long-Memory Time Series Models*. Ph. D. thesis, Southern Methodist University.

Giraitis, L. and R. Leipus (1995). A generalized fractionally differencing approach in long-memory modeling. *Lietuvos Matematikos Rinkinys 35*(1), 65–81.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis 1*, 15–29.

Gray, H. L., N.-F. Zhang, and W. A. Woodward (1989). On generalized fractional processes. *Journal of Time Series Analysis 10*(3), 233–257.

Gray, H. L., N.-F. Zhang, and W. A. Woodward (1994). On generalized fractional processes – a correction. *Journal of Time Series Analysis 15*(5), 561–562.

Hosking, J. R. M. (1981). Fractional differencing. *Biometrika 68*(1), 165–176.

Hosking, J. R. M. (1982). Some models of persistence in time series. In O. D. Anderson (Ed.), *Time Series Analysis: Theory and Practice*, Volume 1, pp. 641–653. Amsterdam: North Holland.

Hosking, J. R. M. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research 20*(12), 1898–1908.

Mallat, S. and Z. Zhang (1993). Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing 41*(12), 3397–3415.

Masry, E. (1993). The wavelet transform of stochastic processes with stationary increments and its application to fractional Brownian motion. *IEEE Transactions on Information Theory 39*(1), 260–264.

McCoy, E. J. and A. T. Walden (1996). Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics 5*(1), 26–56.

Morris, J. M. and R. Peravali (1999). Minimum-bandwidth discrete-time wavelets. *Signal Processing 76*(2), 181–193.

Ooms, M. (1995). Flexible seasonal long memory and economic time series. Technical Report

9515/A, Econometric Institute, Erasmus University.

Percival, D. B. (1992). Simulating Gaussian random processes with specified spectra. *Computing Science and Statistics 24*, 534–538.

Percival, D. B. and A. T. Walden (1999). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press. Forthcoming.

Piessons, R., E. de Doncker-Kapenga, C. W. Überhuber, and D. K. Kahaner (1983). *QUADPACK: A Subroutine Package for Automatic Integration*, Volume 1 of *Springer Series in Computational Mathematics*. Heidelberg: Springer-Verlag.

Rainville, E. D. (1960). *Special Functions*. New York: The Macmillan Company.

Ribeiro, V. J., R. H. Riedi, M. S. Crouse, and R. G. Baraniuk (1999). Simulation of nonGaussian long-range dependent traffic using wavelets. In *ACM SIGMETRICS Conference on the Measurement and Modeling of Computer Systems*. 1-4 May 1999, Atlanta, Georgia.

Tewfik, A. H. and M. Kim (1992). Correlation structure of the discrete wavelet coefficients of fractional Brownian motion. *IEEE Transactions on Information Theory 38*(2), 904–909.

Wickerhauser, M. V. (1994). *Adapted Wavelet Analysis from Theory to Software*. Wellesley, Massachusetts: A K Peters, Ltd.

Woodward, W. A., Q. C. Cheng, and H. L. Gray (1998). A k-factor GARMA long-memory model. *Journal of Time Series Analysis 19*(4), 485–504.

Wornell, G. W. (1993). Wavelet-based representations for the $1/f$ family of fractal processes. *Proceedings of the IEEE 81*(10), 1428–1450.

Wornell, G. W. (1996). *Signal Processing with Fractals: A Wavelet Based Approach*. New Jersey: Prentice Hall.

					DWPT		
Model	N	Hosking	MB(8)	D(16)	LA(16)	MB(16)	MB(24)
$\delta = 0.4$	128	46.44	55.26	54.36	54.57	56.80	54.81
$f_0 = \frac{1}{12}$, $\phi = 0.866$	256	8.83	12.84	10.72	12.98	8.54	8.50
	512	1.55	7.25	7.31	7.86	4.79	5.48
	1024	0.18	7.48	7.19	7.32	4.55	5.38
	2048	0.08	7.39	6.28	6.97	4.67	5.15
$\delta = 0.2$	128	0.0817	0.0872	0.0889	0.0863	0.1002	0.0917
$f_0 = \frac{1}{12}$, $\phi = 0.866$	256	0.0148	0.0184	0.0223	0.0136	0.0132	0.0147
	512	0.0013	0.0121	0.0140	0.0120	0.0063	0.0083
	1024	0.0003	0.0114	0.0103	0.0131	0.0084	0.0067
	2048	0.0002	0.0130	0.0113	0.0109	0.0058	0.0075
$\delta = 0.3$	128	0.144	0.262	0.241	0.206	0.154	0.177
$f_0 = 0.352$, $\phi = -0.6$	256	0.115	0.351	0.371	0.320	0.204	0.198
	512	0.053	0.304	0.256	0.263	0.137	0.149
	1024	0.007	0.306	0.277	0.245	0.115	0.124
	2048	0.001	0.284	0.255	0.253	0.126	0.118
$\delta = 0.3$	128	0.542	0.591	0.631	1.043	0.875	0.615
$f_0 = 0.016$, $\phi = 0.995$	256	0.151	3.010	2.738	3.120	3.313	3.183
	512	0.064	1.686	2.062	1.689	2.367	1.998
	1024	0.056	1.895	2.043	1.544	1.721	1.711
	2048	0.027	1.884	1.970	1.679	1.768	1.747

Table 1: Average squared difference between the mean sample ACVS (1000 simulations) and the true ACVS for lags $0, \ldots, 100$. The Daubechies family of extremal phase compactly supported wavelets is denoted by D(L), the Daubechies family of least asymmetric compactly supported wavelets by LA(L), and the minimum-bandwidth discrete-time wavelets by MB(L) – where L is the number of nonzero coefficients. The basis used for the DWPT method is $\mathcal{B}_7 = \{(7, n) \,|\, n = 0, \ldots, 63\}$ for all wavelet filters.

			DWPT				
Model	N	Hosking	MB(8)	D(16)	LA(16)	MB(16)	MB(24)
$\delta = 0.4$	128	46.44	53.66	54.56	55.49	59.75	53.30
$f_0 = \frac{1}{12}$, $\phi = 0.866$	256	8.83	11.97	10.05	10.58	8.95	10.12
	512	1.55	8.86	6.44	8.55	4.36	5.20
	1024	0.18	7.76	6.78	7.74	4.45	4.71
	2048	0.08	7.80	7.36	7.17	4.46	5.23
$\delta = 0.2$	128	0.0817	0.1559	0.1203	0.1258	0.1304	0.1340
$f_0 = \frac{1}{12}$, $\phi = 0.866$	256	0.0148	0.0715	0.0605	0.0549	0.0751	0.0685
	512	0.0013	0.0682	0.0551	0.0560	0.0571	0.0677
	1024	0.0003	0.0662	0.0547	0.0524	0.0620	0.0667
	2048	0.0002	0.0670	0.0559	0.0536	0.0628	0.0669
$\delta = 0.3$	128	0.144	0.264	0.205	0.237	0.166	0.154
$f_0 = 0.352$, $\phi = -0.6$	256	0.115	0.363	0.391	0.360	0.252	0.139
	512	0.053	0.329	0.300	0.277	0.158	0.154
	1024	0.007	0.297	0.292	0.248	0.125	0.117
	2048	0.001	0.304	0.249	0.253	0.137	0.128
$\delta = 0.3$	128	0.542	0.942	0.831	0.627	0.725	1.083
$f_0 = 0.016$, $\phi = 0.995$	256	0.151	2.917	3.296	2.406	3.530	3.606
	512	0.064	1.988	2.303	1.855	2.170	2.226
	1024	0.056	1.812	1.893	1.719	1.900	1.868
	2048	0.027	1.870	1.788	1.806	1.687	1.759

Table 2: Average squared difference between the mean sample ACVS (1000 simulations) and the true ACVS for lags $0, \ldots, 100$. The DWPT method uses an adaptive orthonormal basis. It depends on the Gegenbauer frequency and length of the wavelet filter, for the first SPP see Figure 4 for the basis functions $\mathcal{B}_L$, $L \in \{8, 16\}$.

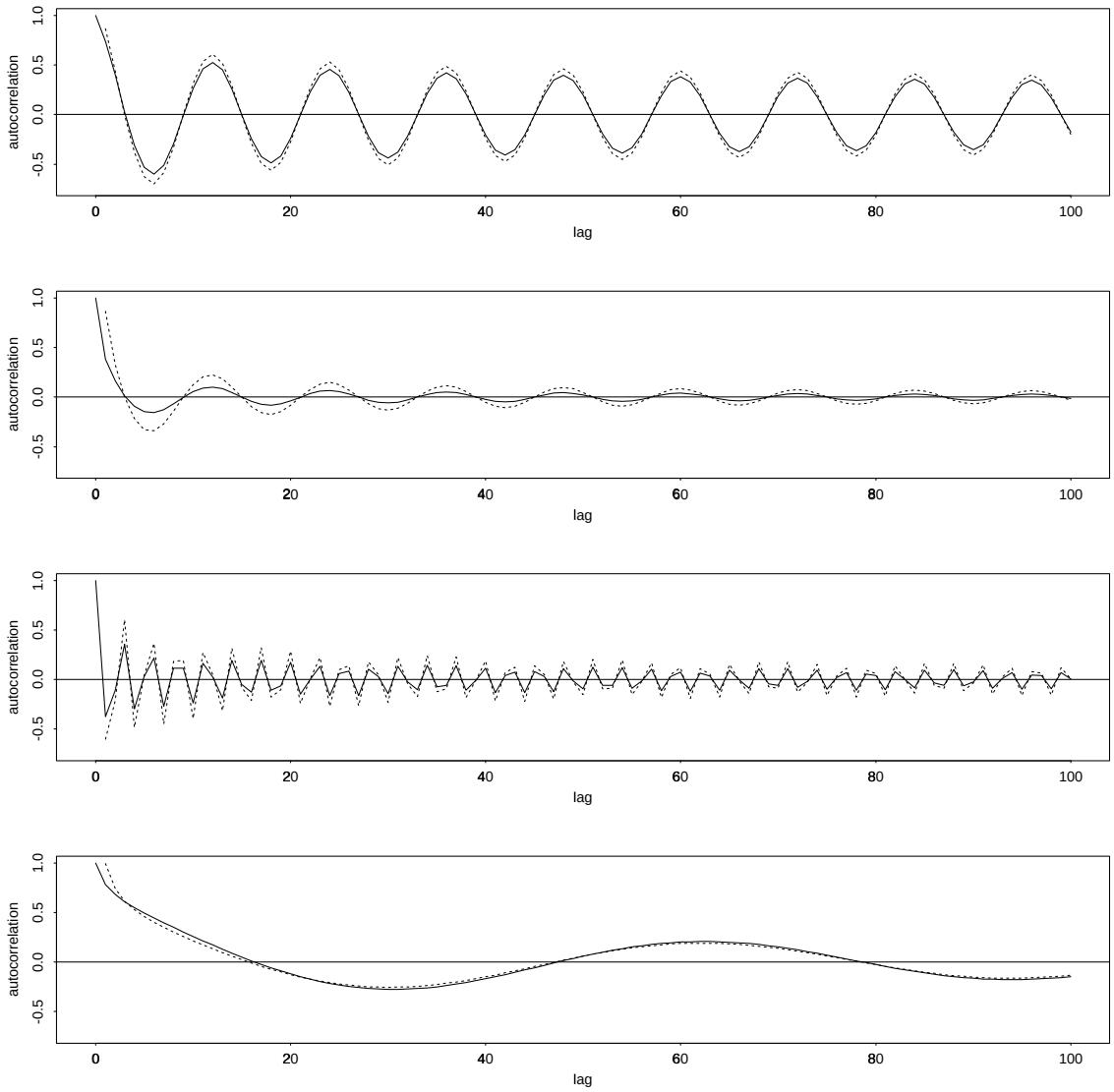
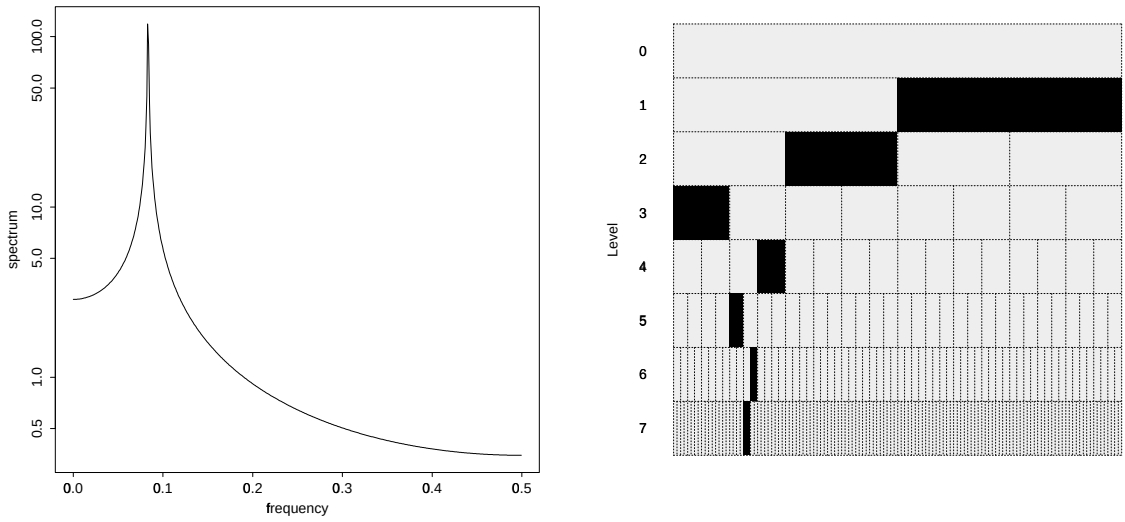


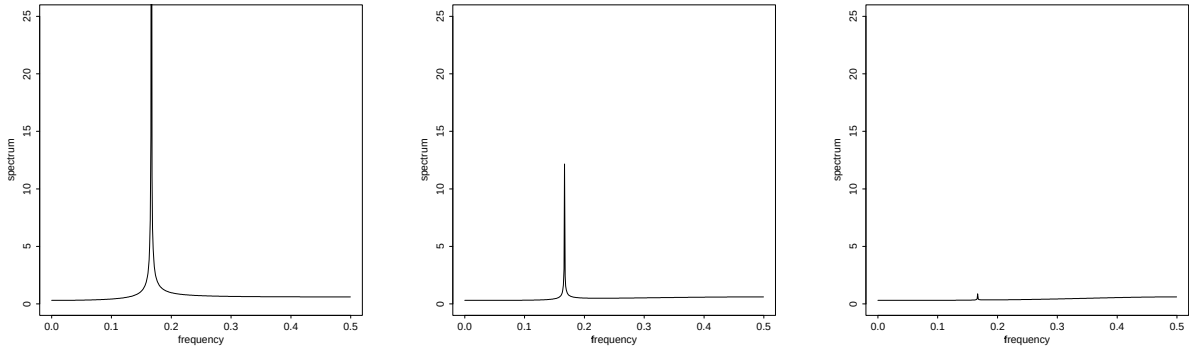
Figure 1: Autocorrelation sequences $\{\rho_\tau\}$ for particular SPPs computed via numeric integration (solid line) and asymptotic approximation (dotted line). From top to bottom the parameters which define the processes are $(\delta = 0.4, f_0 = \frac{1}{12}, \phi = 0.866)$, $(\delta = 0.2, f_0 = \frac{1}{12}, \phi = 0.866)$, $(\delta = 0.3, f_0 = 0.352, \phi = -0.6)$ and $(\delta = 0.3, f_0 = 0.016, \phi = 0.995)$. These correspond to the processes displayed in Figures 8-11 from Anděl (1986).



(a) Spectral density function

(b) Wavelet packet table

Figure 2: Basis selection for an SPP. On the left, the SDF for the SPP ($f_0 = \frac{1}{12}, \delta = 0.4$) and, on the right, its ideal basis ($J = 7$) consisting of – from left to right – $\mathbf{W}_{3,0}$, $\mathbf{W}_{5,4}$, $\mathbf{W}_{7,20}$, $\mathbf{W}_{7,21}$, $\mathbf{W}_{6,11}$, $\mathbf{W}_{4,3}$, $\mathbf{W}_{2,1}$, and $\mathbf{W}_{1,1}$.



(a) $L = 2$

(b) $L = 4$

(c) $L = 8$

Figure 3: Spectrum of $\mathbf{W}_{1,1}$ for the SPP ($f_0 = \frac{1}{12}, \delta = 0.4$) using the Daubechies families of wavelet filters of lengths $L \in \{2, 4, 8\}$. The spectrum is plotted on a linear scale.

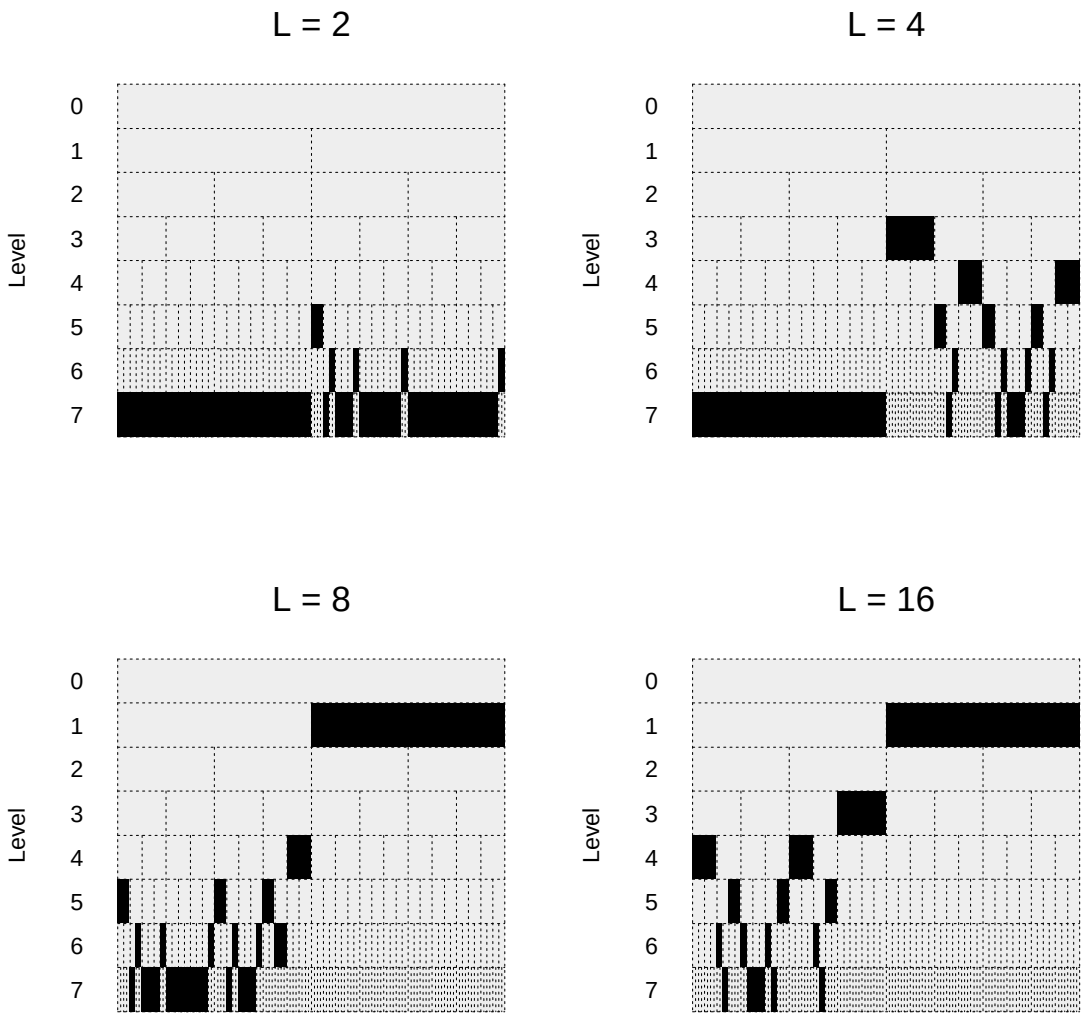


Figure 4: 'Best' orthonormal basis $\mathcal{B}_L$ for the SPP ($f_0 = \frac{1}{12}$, $\delta = 0.4$) based on the Daubechies family of wavelet filters with length $L \in \{2, 4, 8, 16\}$.

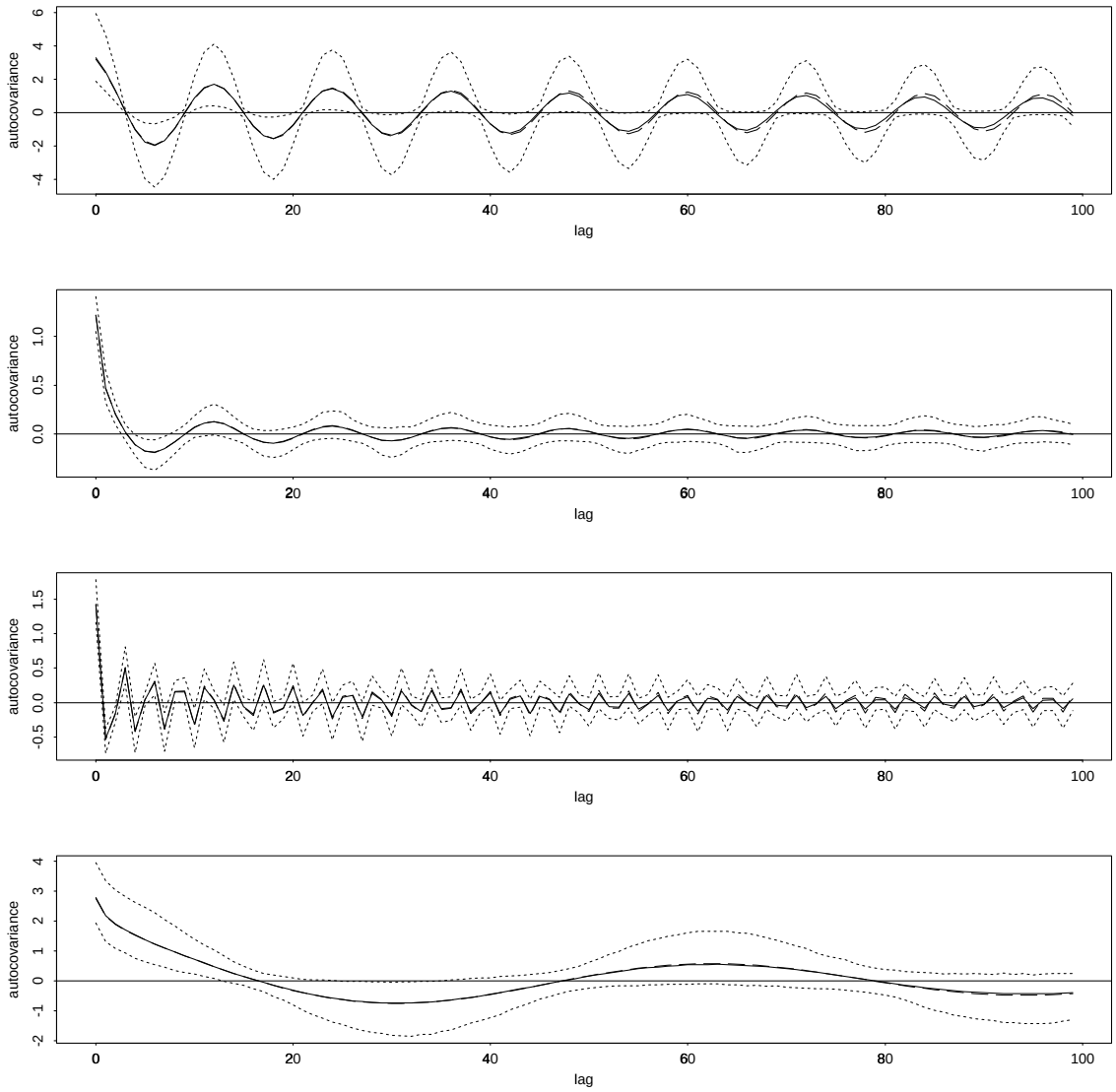


Figure 5: Sample autocovariance sequences (ACVS) for the Hosking method averaged over 500 simulations. Four SPPs ($N = 512$) are displayed with parameters – from top to bottom – ($\delta = 0.4, f_0 = \frac{1}{12}, \phi = 0.866$), ($\delta = 0.2, f_0 = \frac{1}{12}, \phi = 0.866$), ($\delta = 0.3, f_0 = 0.352, \phi = -0.6$) and ($\delta = 0.3, f_0 = 0.016, \phi = 0.995$). The true ACVS is given by the dashed line and the average sample ACVS is given by the solid line. Also shown are the 5% and 95% points (dotted lines).

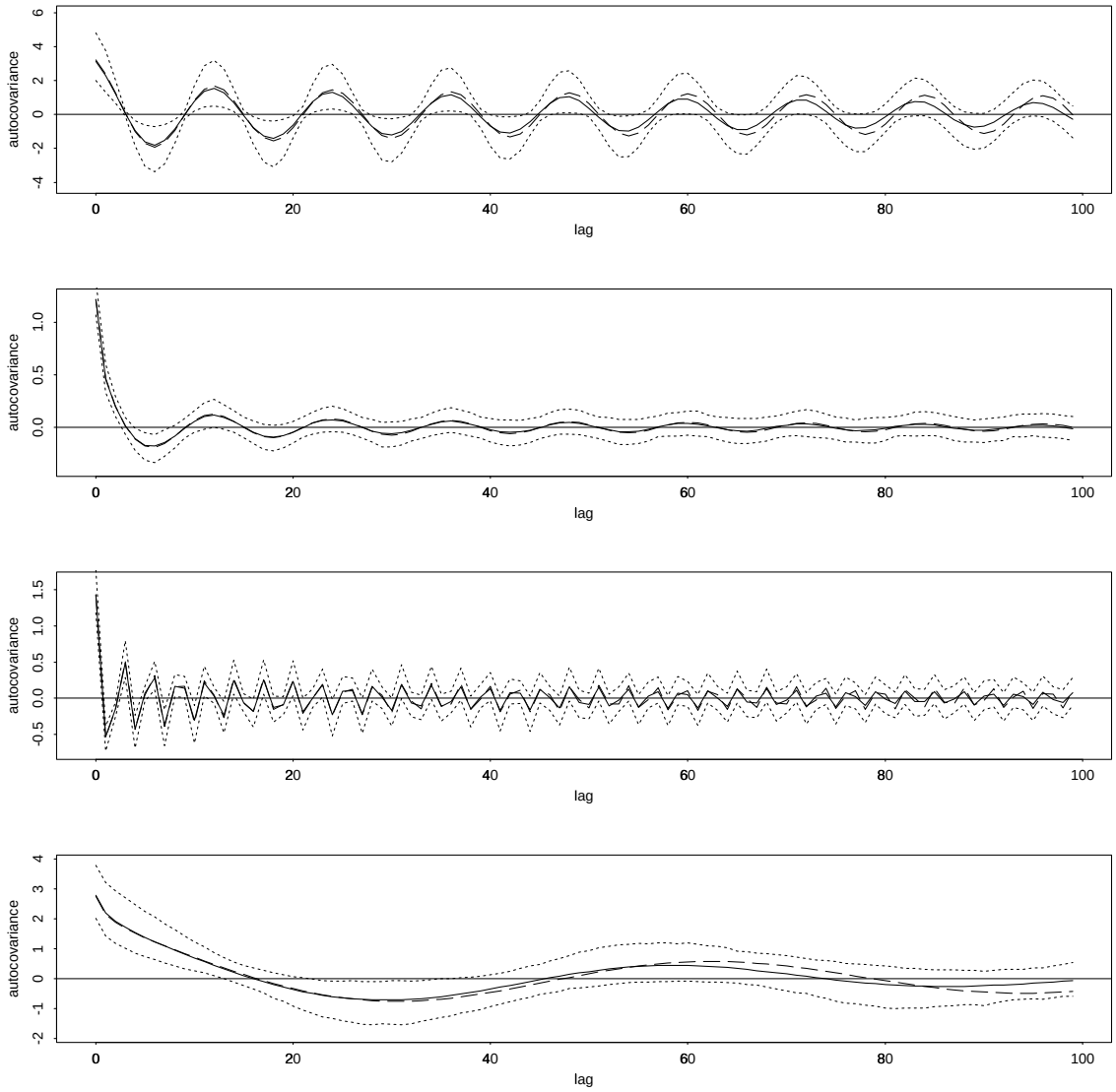


Figure 6: Sample autocovariance sequence (ACVS) for the DWPT method averaged over 500 simulations using the MB(16) wavelet filter and $\mathcal{B}_7 \equiv \{(7,n)\,|\,n=0,\ldots,63\}$ basis. The four SPPs are identical to those in Figure 5. The true ACVS is given by the dashed line and the average sample ACVS is given by the solid line. Also shown are the 5% and 95% points (dotted lines).

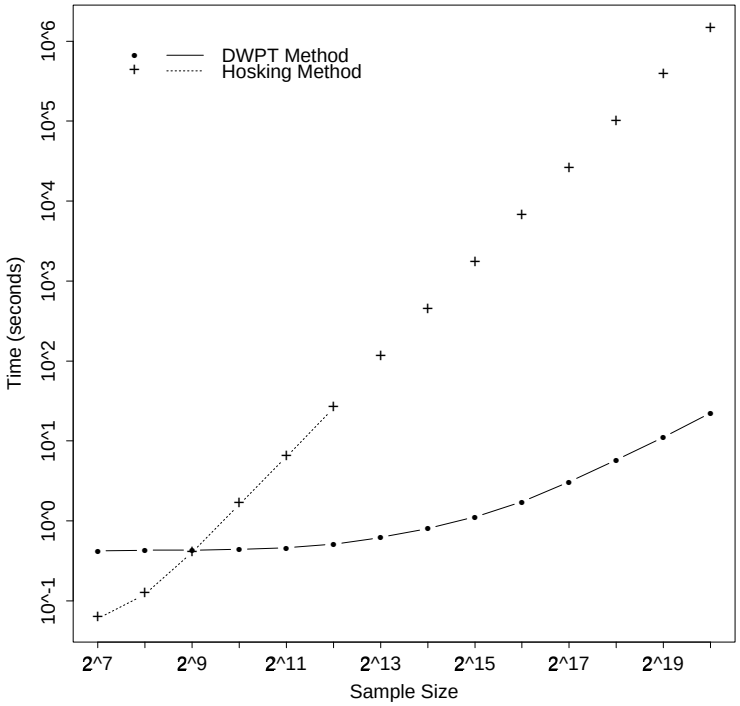


Figure 7: Computation time (in seconds) for the DWPT and Hosking methods to simulate one realization of an SPP. All calculations for the DWPT were implemented in S-Plus. For the Hosking method, sample sizes $N = 2^7, \ldots, 2^{12}$ were computed in S-Plus, while larger sample sizes could not be evaluated. Subsequent computation times for the Hosking method were extrapolated from the previous values.